AD A013212

NRL Report 7900

Cryptographic Digital Communication

DON J. TORRIERI

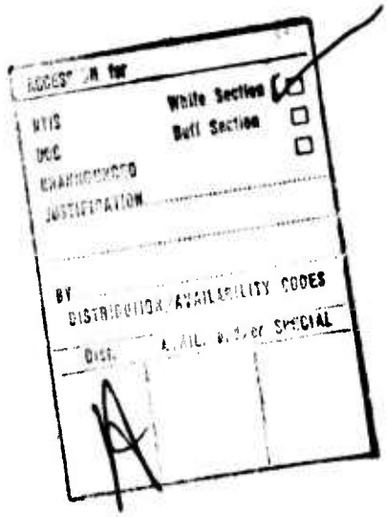
Advanced Projects Office

July 18, 1975

DDC
RECEIVED
AUG 8 1975
D

NAVAL RESEARCH LABORATORY
Washington, D.C.

Approved for public release; distribution unlimited.


ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER: NRL ~~Report~~ 7900	2 GOVT ACCESSION NO.
3 RECIPIENT'S CATALOG NUMBER	
4 TITLE (and Subtitle): CRYPTOGRAPHIC DIGITAL COMMUNICATION	5 TYPE OF REPORT & PERIOD COVERED: A final report on one phase of the NRL problem.
	6 PERFORMING ORG. REPORT NUMBER
7 AUTHOR(s): Don J. Torrieri	8 CONTRACT OR GRANT NUMBER(s): NRL Problem R06-55 ~~Project No. ND 62.00.0~~
9 PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Research Laboratory, Washington, D.C. 20375	10 PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11 CONTROLLING OFFICE NAME AND ADDRESS: Department of the Navy, Naval Electronic Systems Command, Washington, D.C. 20375	12 REPORT DATE: July 1975
	13 NUMBER OF PAGES: 24
14 MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 16 NRL-R06-55	15 SECURITY CLASS. (of this report): Unclassified
	15a DECLASSIFICATION/DOWNGRADING SCHEDULE

16 DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DDC RECEIVED AUG 8 1975 D

17 DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18 SUPPLEMENTARY NOTES

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cryptography
Digital communication
Cryptographic digital communication
Enciphering systems
Stream ciphers
Block ciphers
Word error rate
Bit error rate
Error-rate degradation

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

The fundamentals of cryptographic digital systems are reviewed. The basic types of ciphers are defined and discussed. The remainder of the report examines the word and bit error rates in digital communication systems with block or stream ciphers. Upper bounds and ensemble averages of the error rates are obtained for both ciphers. The effect of modulation type on the degradation in communication system performance caused by the addition of cryptography is analyzed. A comparison is made between block and stream ciphers with respect to their effects on degradation.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251 150

CONTENTS

PRECEDING PAGE BLANK-NOT FILMED

CRYPTOGRAPHIC DIGITAL COMMUNICATION

INTRODUCTION

There are many ways of safeguarding the transmission of secret information. Cryptography is employed when unauthorized personnel have the technical capability of intercepting and correctly interpreting a secret message. A cryptographic analog communication system, such as the interchanging of frequencies to disguise a voice message, usually requires expensive and complicated instrumentation. Due to the availability of the digital computer, cryptographic digital communication systems are more readily automated.

Cryptographic digital communication is accomplished in two ways. Coding consists of the substitution of groups of bits of variable length for plaintext groups of variable length. Encipherment consists of the substitution of fixed-length groups of bits for fixed-length plaintext groups. In general, coding is too slow for high-density data transmission. Another disadvantage is the technical difficulty entailed in the frequent code changes necessary for secrecy. For these reasons, enciphering systems, which provide high-speed capabilities and are easily modified, are used in most practical cryptographic digital communications.

There are two basic types of encipherment — the stream cipher and the block cipher. The stream cipher is bit-by-bit encipherment which results when a binary symbol is added, modulo two, to each bit of plaintext. The complete set of binary symbols or the rule for generating it is called the key. Deciphering is accomplished by adding the key to the corresponding enciphered bit. The more random the key, the more difficult it is for a cryptanalyst to decipher an intercepted cryptogram. Algorithms exist for generating long pseudorandom keys from two or more short streams of digits. However, an algorithm implies a degree of regularity, which enhances the possiblity that an unauthorized cryptanalyst may discern the pattern and duplicate the key generator.

A block cipher is defined as the conversion of m plain bits simultaneously into n enciphered bits. Each of the enciphered bits is a function of all of the plain bits. For unambiguous deciphering, it is necessary that $n \geq m$. For ease of automation, it is preferable that $n = m$. Since knowledge of the conversion of one block of bits reveals little or nothing about the conversion of another block, the block cipher can be made secure by employing large values of n. A practical difficulty is the large number of wires required in the implementation of such a cipher. One might try to circumvent this problem by employing a block cipher which merely transposes the plain bits. However, the simplicity of form of such an enciphering system makes it vulnerable.

Note: Manuscript submitted April 16, 1975.

Many automatic electronic cryptographic systems use a stream cipher which incorporates some of the useful aspects of the block cipher. The technique is to use a pseudo-random key which is a function of the plaintext itself. Thus each enciphered bit is a function of many preceding plain bits. A drawback of this system, which we shall call a data-keyed cipher, is that a single erroneous bit entering the deciphering system causes many additional bit errors down the line. In stream ciphers with data-independent keys, a single error is confined to a single position; in a block cipher, each error may affect any of the other bits of the block.

To consolidate the understanding of the definitions, they shall be put in a more mathematical form. A block cipher is defined to be a rule for associating with each block $(x_i, x_{i+1}, \ldots, x_{i+n})$ of plaintext, a block $(y_i, y_{i+1}, \ldots, y_{i+n})$ of cipher text. Thus we can write

$$y_k = f_k(x_i, x_{i+1}, \ldots, x_{i+n}), \quad i \leq k \leq i + n,$$

where the f_k are functions. A stream cipher is defined to be a rule for associating with each stream $(x_1, x_2, \ldots, x_i, x_{i+1}, \ldots)$ of plaintext, a stream $(y_1, y_2, \ldots, y_i, y_{i+1}, \ldots)$ of cipher text, subject to the restriction

$$y_k = \begin{cases} f_k(x_{k-n+1}, x_{k-n+2}, \ldots, x_k), & k > n; \\ f_k(x_1, x_2, \ldots, x_k), & k \leq n. \end{cases}$$

In addition, y_k is often a function of certain initial conditions in the enciphering and deciphering systems.

Most modern cryptographic systems fit into these two broad categories or represent a hybrid of these two ciphers. For example, the Vernam or one-time system is a stream cipher with a data-independent key; thus, $y_k = f_k(x_k)$.

A special case which aids in the intuitive understanding of the preceding ideas is the linear data-keyed cipher. Figure 1 shows an implementation of a linear data-keyed cipher in which a four-stage shift register of type D flip-flops and exclusive OR gates are used. For each distinct setting of the switches, there is a different enciphered output stream. The corresponding deciphering system is shown in Fig. 2. The extra flip-flop is included for synchronization purposes. The switches must be set in the same manner as those of the enciphering system. A proof of this statement shall now be given.

We define $p(t)$ as the input sequence of plain bits into the enciphering system and $c(t)$ as the corresponding output. Similarly, $c_1(t)$ is the input sequence of enciphered bits into the deciphering system, and $p_1(t)$ is the corresponding output. We define the operation "+" as modulo-two addition. The multiplication is defined as usual.

The operation D is defined by $Dp(t)=p(t - t_0)$, where t_0 is defined such that $t - t_0$ is the time of the clock pulse immediately preceding the time t.

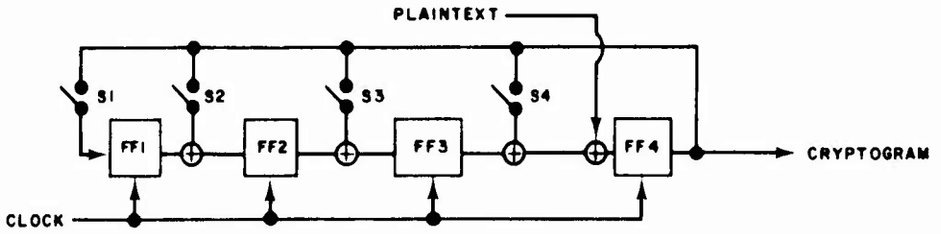


Fig. 1 — Linear data-keyed enciphering system

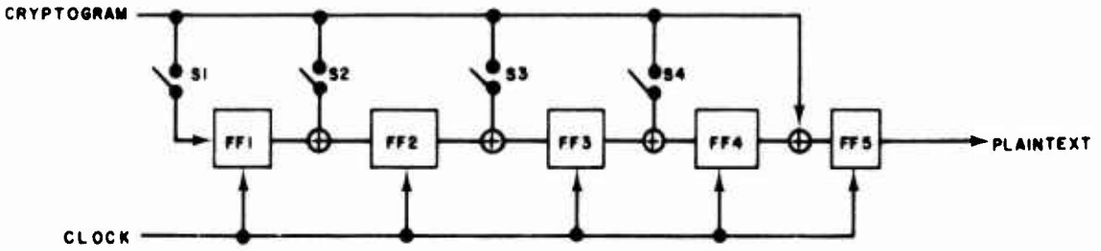


Fig. 2 — Linear data-keyed deciphering system

The discrete variables s_i may take the values 0 or 1, depending on whether the corresponding switch in the enciphering system is open or closed, respectively. The discrete variables s_i' refer to the deciphering system and are defined analogously. With the preceding definitions and the system of Fig. 1, we observe that during steady-state operation,

$$c(t)=Dp(t) + s_4Dc(t) + s_3D^2c(t) + s_2D^3c(t) + s_1D^4c(t). \tag{1}$$

Looking at Fig. 2, we can write

$$p_1(t) = Dc_1(t) + s_4'D^2c_1(t) + s_3'D^3c_1(t) + s_2'D^4c_1(t) + s_1'D^5c_1(t). \tag{2}$$

In modulo-two arithmetic, $a + b = c$ implies $a = b + c$. Using this simple fact, Eq. (1) yields

$$Dp(t) = c(t) + s_4Dc(t) + s_3D^2c(t) + s_2D^3c(t) + s_1D^4c(t). \tag{3}$$

We observe that $c_1(t) = c(t - \tau)$, where τ is the delay due to transmission. Suppose s_i' $= s_i$ for $i = 1, 2, 3, 4$. A comparison of Eqs. (2) and (3) then indicates that

$$p_1(t) = D^2 p(t - \tau). \tag{4}$$

Thus the output of the deciphering system is a delayed version of the input to the enciphering system. The proof for the general system of n shift-register stages is analogous.

In the absence of an input, the system of Fig. 1 behaves as a pseudorandom word generator. The maximum length of the output sequence before pattern repetition is 2^n - 1 bits, where n is the total number of functioning shift-register stages. The maximum-length sequence will occur only for certain switch settings and only if the initial flip-flop states are not all zero. For example, in Fig. 1, switch $S1$ must be closed if a pseudorandom sequence of length 15 is to be generated. If $S1$ is open, the maximum possible length is 7. If $n = 20$, a pseudorandom sequence of over a million bits in length may be generated. It would seem that enciphered bits produced by such a system would be undecipherable with less than $2^n - 1$ intercepted bits; cryptanalysis would be hopeless if $n > 20$. However, we shall show that the key can be broken with as few as $2n$ bits.

Consider the discrete times t_i, where $t_{i+1} = t_i + T$, and T is the clock (bit) period. In the general case, we have the following steady-state relations analogous to Eq. (3):

$$Dp(t_i) = c(t_i) + s_n Dc(t_i) + \ldots + s_1 D^n c(t_i),$$

$$i = 1, 2, \ldots, n. \tag{5}$$

Since $Dc(t_{i+1}) = c(t_i)$, the n equations represented above contain the n unknown values of s_i and $2n$ values of $c(t)$. It follows that it is possible, under the appropriate conditons and with knowledge of the $2n$ values of $c(t)$ and the n values of $Dp(t_i)$, to solve the system of equations for the s_i.

As an example, consider the case where $n = 4$. Suppose we acquire the following sequences of plaintext and enciphered bits:

$c(t_i)$: 1 0 0 1 0 0 1 1 1

$Dp(t_i)$: 1 0 1 0 1 0 1 0 1

The first four values of $Dp(t_i)$ do not help us, since we cannot construct all the terms on the right side of Eq. (5). As a matter of fact, Eq. (5) may not be valid for the first four

values of $Dp(t_i)$ since we have not been told whether the steady state has been reached If the clock of the enciphering system has just started at t_1, then the four values are dependent on the initial states of the flip-flops. From the second set of $n = 4$ observations, Eq. (5) yields

$$1 = s_4 + s_1, \tag{6}$$

$$0 = s_3, \tag{7}$$

$$1 = 1 + s_2, \tag{8}$$

and

$$0 = 1 + s_4 + s_1, \tag{9}$$

which imply that $s_2 = s_3 = 0$, but do not tell us uniquely the values of s_1 and s_4. If we use the final observation, we obtain

$$1 = 1 + s_4 + s_3, \tag{10}$$

which now allows us to assert that $s_4 = 0$ and $s_1 = 1$. Note that $n + 1 = 5$ known plain bits and $2n + 1 = 9$ enciphered bits were used. However, if we had originally used Eqs. (7) — (10) instead of Eqs. (6) — (9), we could have obtained the solution with $n = 4$ known plain bits and $2n = 8$ enciphered bits. Once the switch settings have been determined, it is easy to solve for the initial states.

If the switch $S1$ is open, the first flip-flop is nonfunctional, and we have an enciphering system with only three shift-register stages. However, the cryptanalyst usually does not know a priori the number of shift-register stages. Consequently, he must allow for the largest number of stages possible while attempting to break the key.

There are certain bizarre circumstances under which the key cannot be broken, despite an indefinitely long, known set of plain and enciphered bits. For example, suppose we have the periodic patterns

$$c(t_i): \ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ \ldots\ 0\ 0\ 1\ 1$$

and

$$Dp(t_i): \ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \ldots\ 0\ 0\ 0\ 0.$$

It is readily verified that there are two possible solutions, no matter how many of these patterns are observed.

Since it readily can be cracked under certain circumstances, the linear data-keyed cipher is not a very practical system for high-security purposes. It can be reasonably effective for infrequent, low-security operations if the number of stages is large and if the user is careful not to use plaintexts of many consecutive zeros or ones, too systematic a formatting of frames, or indications of where words start and end. For high-security purposes, nonlinear systems based on operations other than modulo-two arithmetic can be designed to make code breaking extremely complicated and expensive. A block diagram of a general data-keyed enciphering or deciphering system is shown in Fig. 3.

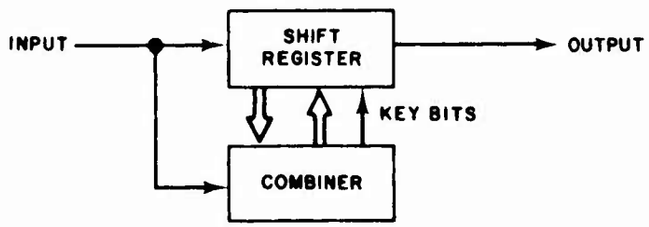


Fig. 3 — General data-keyed enciphering or deciphering system

In any digital communication system, the transmitted bits and words have certain error rates. Except for stream ciphers with data-independent keys, encipherment causes these error rates to increase if other system parameters remain unchanged. In block ciphers, each deciphered bit is a function of all the transmitted enciphered bits in the corresponding block. Therefore a single erroneous received bit is practically certain to cause many erroneous deciphered bits. For the data-keyed system of Fig. 3, the degradation is due to the presence of the shift register. A received bit error due to random noise is carried through the shift register, causing additional bit errors down the line. We shall obtain quantitative measures of the degradation for general stream and block ciphers.

It can be verified easily that the roles of Figs. 1 and 2 can be interchanged; that is, the system of Fig. 2 could serve as an enciphering system with the system of Fig. 1 as the corresponding deciphering system. However, this choice is not a good one for a practical communication network, since a single bit error at the input of Fig. 1 will cause an indefinite number of further errors at the output. In the original configuration, only four output bits at most are affected by a single input bit error at the deciphering system.

ERROR-RATE BOUNDS FOR STREAM CIPHERS

We shall designate by P_b the probability of bit error for an unenciphered communication system. We shall assume that the bit errors resulting from transmission occur independently of each other. It follows that the word error rate is

$$P_w = 1 - (1 - P_b)^k , \tag{11}$$

where k denotes the number of bits per word. We now investigate what happens when a stream cipher is added to the communication system.

Suppose an enciphered bit is erroneously received as a result of random noise or other interference. As the erroneous bit proceeds through the deciphering system, each of n consecutive output bits will be affected. We define a train to be this set of n consecutive bits emerging from the deciphering system. For a stream cipher with a data-independent key, $n = 1$. For a data-keyed cipher, $n > 1$.

The k bits of an enciphered word entering the deciphering system shall be referred to as the input word. The corresponding k plain bits emerging from the deciphering system shall be designated the output word. The probability of a word error, P_{cw}, is defined to be the probability of one or more erroneous bits in the output word. We shall say that a train is of external origin with respect to an output word if the first bit of the train occurs before the first bit of the word. The joint probability of a word error and a train of external origin extending into the word is denoted by $P(w, t)$. If no train of external origin extends into the word, the conditional probability of word error is denoted by $P(w|\bar{t})$. The probability that a train of external origin does not extend into a word is denoted by $P(\bar{t})$. With these definitions and notation, we now derive a decomposition which will be useful in our analysis of stream-cipher error rates.

From the theorem of total probability,

$$P_{cw} = P(w, t) + P(w|\bar{t})P(\bar{t}). \tag{12}$$

A train will extend into an output word if, and only if, one of the $n - 1$ input bits immediately preceding the corresponding input word is in error due to random noise. Thus, assuming bit errors are independent,

$$P(\bar{t}) = (1 - P_b)^{n-1}. \tag{13}$$

When no train is present, an error in one of the bits of the input word causes an error in the corresponding bit of the output word. Thus $P(w|\bar{t})$ is the same as the probability of a word error for plaintext; that is,

$$P(w|\bar{t}) = 1 - (1 - P_b)^k. \tag{14}$$

To determine $P(w, t)$, additional notation must be introduced. If i bits of a train of external origin extend into a word, we denote this condition by the symbols $tb = i$. For example, $P(tb = i)$ denotes the probability that a word contains i externally generated train bits. Since $P(w, t|tb = i) = P(w|tb = i)$, we can write

$$P(w, t) = P(w|tb = k)P(tb = k) + \sum_{i=1}^{k-1} P(w|tb = i)P(tb = i). \tag{15}$$

If at least one of the $n - k$ bits preceding the corresponding input word is in error and $n > k$, it is clear that $tb = k$. Thus

$$P(tb = k) = \begin{cases} 1 - (1 - P_b)^{n-k}, & n > k; \\ 0, & n \leqslant k. \end{cases} \tag{16}$$

For $tb = i$, where $1 \leqslant i < k$, it is necessary that there be an error precisely $n - i$ bits prior to the word bu no erroneous bits among the next $n - i - 1$ bits. Therefore, for $1 \leqslant i < k$,

$$P(tb = i) = \begin{cases} P_b(1 - P_b)^{n-i-1}, & n > i; \\ 0, & n \leqslant i. \end{cases} \tag{17}$$

Substitution of Eqs. (13) through (17) into Eq. (12) yields the decomposition

$$P_{cw} = P(w|tb = k)\left[1 - (1 - P_b)^{n-k}\right] u(n - k) + \sum_{i=1}^{min(k-1,n-1)} P(w|tb = i)P_b(1 - P_b)^{n-i-1} + \left[1 - (1 - P_b)^k\right](1 - P_b)^{n-1}, \tag{18}$$

where $u(n - k)$ is a step function, that is, $u(n - k)$ is 0 for $n < k$ and is 1 for $n \geqslant k$. Note that in the summation term, i extends to the least of the two integers $k - 1$ and $n - 1$.

To evaluate the decomposition, the exact configuration of the cryptographic system has to be specified. However, a tight upper bound can be obtained by simply observing that $P(w|tb = k)$ and $P(w|tb = i)$ must be less than unity. Therefore

$$P_{cw} \leqslant \left[1 - (1 - P_b)^{n-k}\right] u(n - k) + \sum_{i=1}^{min(k-1,\, n-1)} P_b(1 - P_b)^{n-i-1} + \left[1 - (1 - P_b)^k\right](1 - P_b)^{n-1} \tag{19}$$

After some algebraic simplification, Eq. (19) reduces to the compact expression

$$P_{cw} \leqslant 1 - (1 - P_b)^{n+k-1}. \tag{20}$$

We shall now show that there is a simpler bound

$$P_{cw} \leqslant (n + k - 1)P_b. \tag{21}$$

Consider the function of P_b defined by

$$y = (n + k - 1)P_b - 1 + (1 - P_b)^{n+k-1}. \tag{22}$$

Clearly y is zero at $P_b = 0$. Since $n + k \geqslant 2$, y has a nonnegative derivative for all P_b such that $0 \leqslant P_b \leqslant 1$. Thus for all possible P_b, $y \geqslant 0$. We conclude that

$$(n + k - 1)\, P_b \geqslant 1 - (1 - P_b)^{n+k-1}. \tag{23}$$

Combining Eqs. (20) and (23) yields Eq. (21).

Using $k = 1$ in Eq. (21), we obtain the companion inequality

$$P_{cb} \leqslant nP_b. \tag{24}$$

A binomial expansion indicates that the bound of Eq. (21) is almost as tight as the bound of Eq. (20) if

$$P_b << 2(n + k - 2)^{-1}, \quad n + k > 2. \tag{25}$$

ENSEMBLE-AVERAGE ERROR RATES FOR STREAM CIPHERS

A second measure of error-rate performance is obtained by considering ensembles of stream ciphers characterized by a specific value of the parameter n. In what follows, we indicate an ensemble average by a bar over the P. Let the symbol X denote the ensemble-average probability that a bit which is part of a train of external origin is in the correct state. Before deriving an expression for $\overline{P}_{cw}$ we shall first investigate what value X might have.

For linear systems, X is one-half, independent of the input word and the other output bits. This statement is also true if a bit is simultaneously part of two or more trains.

To see the truth of this assertion, consider the linear system of Fig. 2. Suppose that after n correct input bits, an erroneous input bit is received. The corresponding output bit is then in error, and a train is started. Over the ensemble of deciphering systems of the form of Fig. 2, it is equally likely that $S4$ will be open or closed. If $S4$ is closed and the next input bit is correct, it is seen that the next output bit is in error. Similarly, if $S4$ is open and the next input bit is in error, the next output bit is in error. Thus if the next input bit has an error probability P_b, the error probability of the next output bit is $(1/2)(1 - P_b) + (1/2)P_b = 1/2$. Continuing this reasoning leads to the conclusion that $X = 1/2$.

It is believed that X is one-half with respect to the ensemble of all possible stream ciphers, independent of the input word and the other output bits. Referring to Fig. 3, notice that over the ensemble an enciphered input can be applied simultaneously to any number of the shift-register stages and combiner elements. Also, any number of the shift-register outputs can feed the combiner. Because of the nonlinear operation of the combiner, an error in one or more of the bits feeding it may or may not produce an erroneous key bit. Thus in the ensemble there are deciphering systems in which a single erroneous input bit causes several bad bits to be fed into the combiner during most of the key production, and the nonlinear operation causes the subsequent bit error rate to be greater than one-half. Clearly, in the ensemble there are other systems about which the opposite is true.

Although X is one-half for the complete ensemble of all possible stream ciphers, it is possible that for a subset of nonlinear stream ciphers, X is different than one-half with respect to the restricted ensemble. However, the most important practical stream cipher subset is the subset of secure ciphers, that is, those systems for which cryptanalysis is very difficult. Setting X equal to one-half for this subset is an excellent approximation.

When $k = 1$, $\overline{P}(w|tb=k) = 1 - X$. Thus it follows from Eq. (18) that

$$\overline{P}_{cb} = (1 - X)\left[1 - (1 - P_b)^{n-1}\right] + P_b(1 - P_b)^{n-1}. \tag{26}$$

In this equation we have kept the unspecified parameter X because its retention does not complicate the expression significantly. However, for the reasons mentioned and to facilitate the derivation, we shall always assume $X = 1/2$ in determining the ensemble-average word error rate.

We denote the condition that one or more of the first i bits of an input word is in error by the symbol α and the absence of the condition by $\overline{\alpha}$. Using the theorem of total probability, we can write

$$\overline{P}\,(w|tb = i) = \overline{P}\,(w,\alpha|tb = i) + \overline{P}\,(w,\overline{\alpha}|tb = i). \tag{27}$$

If $tb = i$, the ensemble-average probability of no error in the first i output bits is $(1/2)^i$, independent of the input bits and the other output bits. If α is false and $tb = i$, the last $k - i$ output bits are not part of a train generated by the first i bits. Consequently the first error in the last $k - i$ input bits is added to a good key bit. Therefore, the probability

of no error in the last $k - i$ output bits is equal to the probability of no error in the corresponding input bits. We conclude that

$$\overline{P}(w|tb=i, \overline{\alpha}) = 1 - 2^{-i}(1 - P_b)^{k-i}. \tag{28}$$

From the independence of bit errors, we have

$$P(\overline{\alpha}|tb = i) = (1 - P_b)^i. \tag{29}$$

From the definition of a conditional probability and Eqs. (28) and (29),

$$\overline{P}(w,\overline{\alpha}|tb = i) = (1 - P_b)^i - 2^{-i}(1 - P_b)^k. \tag{30}$$

In almost all practical systems, we have $n \geqslant k$. Thus, deferring consideration of the more complicated general case until later, we assume that $n \geqslant k$ and determine $\overline{P}(w,\alpha|tb = i)$ in a manner similar to the derivation of Eq. (30). Clearly

$$P(\alpha|tb = i) = 1 - (1 - P_b)^i. \tag{31}$$

If α is true, $n \geqslant k$, and $tb = i$, then every output bit is part of a train. Consequently the ensemble-average probability of no error for each output bit is 1/2, independent of the other output bits. It follows that

$$\overline{P}(w|tb = i, \alpha) = 1 - 2^{-k}, \quad n \geqslant k. \tag{32}$$

From the definition of a conditional probability and using Eqs. (30), (31), and (32) in Eq. (27), there results

$$\overline{P}(w|tb = i) = 1 - 2^{-i}(1 - P_b)^k - 2^{-k}\left[1 - (1 - P_b)^i\right], \quad n \geqslant k. \tag{33}$$

From this relation or by direct reasoning it follows that for $n > k$,

$$\overline{P}(w|tb = k) = 1 - 2^{-k}. \tag{34}$$

Substitution of Eqs. (33) and (34) into Eq. (18) gives the ensemble-average word error rate. After performing two easy summations and regrouping, we obtain

$$\overline{P}_{cw} = 1 - 2^{-k} + k2^{-k}P_b(1 - P_b)^{n-1} - (1 - P_b)^{n+k-1} + 2^{-k}(1 - P_b)^n - \sum_{i=1}^{k-1} 2^{-i}P_b(1 - P_b)^{n+k-1-i}. \tag{35}$$

Although we shall soon apply it in the present form, this equation can be made slightly more convenient for computation by performing the remaining summation to obtain, for $n \geq k$,

$$\overline{P}_{cw} = \begin{cases} 1 - 2^{-k} + k2^{-k}P_b(1-P_b)^{n-1} - \dfrac{(1-P_b)^n\left[(1-P_b)^k - 2^{-k}\right]}{1-2P_b}, & P_b \neq 1/2; \\ 1 - 2^{-k}, & P_b = 1/2. \end{cases} \tag{36}$$

This formula is still tedious to use in manual computations. Fortunately, a simple asymptotic expression is highly accurate over the usual range of interest. The approximation can be obtained by employing a Taylor-series expansion about the point $P_b = 0$ and dropping the higher order terms. However, the condition for the validity of this procedure is too complicated for quick verification. Consequently, we use an alternative method which yields a simple sufficient condition of validity. Each of the factors in Eq. (35) of the form $(1 - P_b)^m$ is approximated by $1 - mP_b$; a sufficient condition for this approximation is $P_b << 2(m-1)^{-1}$ if $m > 1$. Each factor of the form $P_b(1 - P_b)^m$ is approximated by P_b; a sufficient condition for this approximation is $P_b << m^{-1}$ if $m > 0$. With these approximations and some algebraic simplification, Eq. (35) reduces to

$$\overline{P}_{cw} \approx \left[n + k - 2 - 2^{-k}(n - k - 2)\right]P_b, \quad n \geq k. \tag{37}$$

Combining all the conditions which arise, it is found that the single condition

$$P_b << (n + k - 2)^{-1}, \quad n + k > 2, \tag{38}$$

suffices; that is, Eq. (38) is a sufficient condition for the validity of Eq. (37). Using the same method on Eq. (26), we obtain

$$\overline{P}_{cb} \approx \left[n(1 - X) + X\right]P_b. \tag{39}$$

For later comparison, we note that the asymptotic form of Eq. (11) is

$$P_w \approx kP_b. \tag{40}$$

It is readily verified that Eq. (38) is also a sufficient condition for the validity of Eqs. (39) and (40).

To include the possibility that $n < k$, we must employ more intricate reasoning. Let the symbol $\beta = l$ designate the condition that the last bit error among the first i input bits occurs at input bit l, where $1 \leq l \leq i$. If $\beta = l$, then α is true; thus we make the decomposition

$$\overline{P}(w,\alpha|tb = i) = \sum_{l=1}^{i} \overline{P}(w|tb = i, \beta = l)P(\beta = l|tb = i). \tag{41}$$

Clearly the probability of $\beta = l$, given $tb = i$, is equal to the probability that input bit l is erroneous and input bits $l + 1$ through i are correct. We conclude that

$$P(\beta = l|tb = i) = P_b(1 - P_b)^{i-l}, \quad 1 \leqslant l \leqslant i. \tag{42}$$

When $tb = i$, the probability that the first i output bits are correct has an ensemble average equal to 2^{-i}. The probability that the last $k - i$ output bits are correct depends only on the condition $\beta = l$, which implies that a train of $n + l - i - 1$ bits extends into the final $k - i$ bits. Let w_{k-i} denote an error in a word consisting of $k - i$ output bits. From the previous discussion it follows that

$$\overline{P}(w|tb = i, \beta = l) = 1 - 2^{-i}\left[1 - \overline{P}(w_{k-i}|tb = n + l - i - 1)\right]. \tag{43}$$

Substituting Eqs. (42) and (43) into Eq. (41) gives

$$\overline{P}(w, \alpha|tb = i) = P_b \sum_{l=1}^{i} (1 - P_b)^{i-l}\left[1 - 2^{-i} + 2^{-i}\,\overline{P}(w_{k-i}|tb = n + l - i - 1)\right]. \tag{44}$$

Using Eqs. (44) and (30) in Eq. (27), we obtain

$$\overline{P}(w|tb = i) = 1 - 2^{-i}(1 - P_b)^k - 2^{-i}\left[1 - (1 - P_b)^i\right]$$

$$+ 2^{-i}P_b \sum_{l=1}^{i} (1 - P_b)^{i-l}\,\overline{P}(w_{k-i}|tb = n + l - i - 1). \tag{45}$$

This expression is valid for all n. When $n \geqslant k$, $\overline{P}(w_{k-i}|tb = n + l - i - 1) = 1 - 2^{-(k-i)}$, independent of l. Consequently Eq. (45) reduces to Eq. (33). However, when $n < k$, $\overline{P}(w_{k-i}|tb = n + l - i - 1)$ must be evaluated by the same procedure as that leading to Eq. (45) itself. In general, we have a finite hiearchy of equations, with the number of equations depending on $k - n$. The general ensemble-average cryptographic word-error-rate formula follows on substitution of Eqs. (34) and (45) into Eq. (18).

To obtain an asymptotic expression for $\overline{P}_{cw}$ when $n < k$, we note that the last term in Eq. (45) does not contribute to the final equation even if $\overline{P}(w_{k-i}|tb = n + l - i - 1) = 1$. Then, applying the method described previously to Eqs. (45) and (18), we obtain

$$\overline{P}_{cw} \approx \left[n + k - 2\,(1 - 2^{-n})\right] P_b, \quad n < k, \tag{46}$$

where Eq. (38) provides a sufficient condition.

ERROR-RATE BOUNDS AND ENSEMBLE AVERAGES FOR BLOCK CIPHERS

In the conventional block cipher, a plaintext block of m total bits, comprising an integral number of words of k bits each, is enciphered as a block of n total bits. After transmission and reception, the plaintext block is restored as the output of the deciphering system. Clearly no output words will be in error unless the received enciphered block contains an error in at least one of its n bits. Thus we can write

$$P_{cw} = P(w|be)\left[1 - (1 - P_b)^n\right], \tag{47}$$

where $P(w|be)$ is the probability of an error in an output word, given that there is a block error at the input of the deciphering system. Setting $P(w|be) = 1$ and using Eq. (23), we see that Eq. (47) yields the upper bound given by

$$P_{cw} \leqslant nP_b \tag{48}$$

If $k > 1$, this upper bound is less than the corresponding upper bound for the stream cipher, given by Eq. (21). Since the parameter k does not appear in Eq. (48), the right-hand side provides an upper bound for P_{cb} also. For P_{cb} the upper bound is the same as that indicated in Eq. (24) for the stream cipher.

Usually block ciphers do not involve a size change, that is, $n = m$. We proceed to obtain the ensemble-average cryptographic error rates for this case. Due to the one-to-one correspondence between the enciphered and plaintext blocks, an error in a received enciphered block is certain to cause at least one erroneous bit in the output block. Consequently, over the ensemble of block ciphers there are $2^n - 1$ equally likely output blocks corresponding to an erroneous enciphered block. Consider any fixed bit in these output blocks. In $2^{n-1} - 1$ of the possible output blocks, this bit will be correct, that is, in the same state it would have been if no error had occurred in the enciphered block. We conclude that given a block error, there is an ensemble average probability that a bit is correct equal to $(2^{n-1} - 1)/(2^n - 1)$. Consider a second fixed output bit. Given that there is a block error and that the first fixed output bit is correct, it follows from an extension of the previous reasoning that there is an ensemble-average probability that the second fixed bit is correct equal to $(2^{n-2} - 1)/(2^{n-1} - 1)$. If $x_1, x_2, \ldots x_n$ are events, the probability of all these events can be described as follows:

$$P(x_1, x_2, \ldots, x_n) = P(x_n|x_{n-1}, \ldots, x_1) \ldots P(x_2|x_1)P(x_1). \tag{49}$$

Using Eq. (49) and repeating our analysis for successive output bits, we conclude that

$$\overline{P}(w|be) = 1 - \prod_{i=1}^{k} \frac{2^{n-i} - 1}{2^{n+1-i} - 1}$$

$$= \frac{2^n(1 - 2^{-k})}{2^n - 1}. \tag{50}$$

Combining this relation with Eq. (47), we obtain the ensemble-average cryptographic word error rate for block ciphers

$$\overline{P}_{cw} = (1 - 2^{-n})^{-1} (1 - 2^{-k}) \left[1 - (1 - P_b)^n\right]. \tag{51}$$

The ensemble-average cryptographic bit error rate for block ciphers is

$$\overline{P}_{cb} = 1/2 (1 - 2^{-n})^{-1} \left[1 - (1 - P_b)^n\right]. \tag{52}$$

Under the condition that

$$P_b << 2(n - 1)^{-1}, \tag{53}$$

we obtain the asymptotic formulas

$$\overline{P}_{cb} \approx (1 - 2^{-n})^{-1} \frac{n}{2} P_b \tag{54}$$

and

$$\overline{P}_{cw} \approx (1 - 2^{-n})^{-1}(1 - 2^{-k}) n P_b. \tag{55}$$

Although these formulas hold for all values of n and k, it should be remembered that $n \geqslant 4k$ is usually required to safeguard against the frequency analysis of block patterns. We shall compare the error rates of block and stream ciphers in the next section.

DEGRADATION DUE TO CRYPTOGRAPHY

The bit error rate for ordinary transmission is a function of the modulation system. For most modulation systems, when white Gaussian noise is present, the bit error rate has the functional form specified by

$$P_b = f\left(\frac{E_b}{N_0}\right), \tag{56}$$

where f is a function, N_0 is the noise power spectral density, and E_b is the mean energy for a bit in the one state. If this equation is substituted into Eqs. (26) and (36), or Eqs.

(51) and (52), there result formulas in terms of E_b. By comparing these formulas with Eqs. (11) and (56), we can determine the increase in E_b required to obtain the same error rate from a cryptographic system as the corresponding plaintext system. This increase provides a quantitative measure of cryptographic degradation. Let P'_{cw} denote either $\overline{P}_{cw}$ or the upper bound of P_{cw}. Then the degradation in decibels is defined to be

$$D = 10 \log_{10} \frac{E'_b}{N_0} - 10 \log_{10} \frac{E_b}{N_0} = 10 \log_{10} \frac{E'_b}{E_b}, \tag{57}$$

where E'_b is the energy required to produce a value of P'_{cw} which is equal to the value of P_w when the energy is E_b.

As an example, suppose we wish to calculate the degradation of the ensemble-average bit error rate of a block cipher relative to the plaintext bit error rate. Suppose Eq. (56) is plotted empirically. Then we can also plot Eq. (52). For each value of P_b, we can read a value of E_b/N_0 from the first plot and a value of E'_b/N_0 corresponding to $\overline{P}_{cb} = P_b$ from the second plot. Substitution into Eq. (57) yields D.

Rather than employ the graphical method, it is often convenient to have a simple approximate formula for degradation. To derive such a formula, note that with the help of Eq. (40) all our asymptotic error rate bounds and ensemble averages can be written in the form

$$P'_{cw} = g(n, k) P_w, \tag{58}$$

where $g(n, k)$ is the corresponding function of the parameters n and k. According to the definition of E'_b, it is implicitly related to E_b by

$$P'_{cw}(E'_b) = P_w(E_b). \tag{59}$$

Combining Eqs. (40), (56), (58), and (59), it follows that the degradation can be determined analytically by solving

$$g(n, k) f\left(\frac{E'_b}{N_0}\right) = f\left(\frac{E_b}{N_0}\right). \tag{60}$$

For conventional, ideal, coherent modulation systems, we can write

$$f\left(\frac{E_b}{N_0}\right) \approx \left(\frac{2cE_b}{N_0}\right)^{-1/2} \exp\left(-\frac{cE_b}{2N_0}\right). \tag{61}$$

where c is a constant depending on the modulation type. This relation depends on the asymptotic approximation

$$erfc\ (x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{x^2}{2}\right) dx \approx \frac{1}{\sqrt{2\pi}\ x} \exp\left(-\frac{x^2}{2}\right), \tag{62}$$

which can be employed with negligible error when $cE_b/N_0 > 10$. For conventional, ideal, noncoherent modulation systems, we can write

$$f\left(\frac{E_b}{N_0}\right) = \frac{1}{2} \exp\left(-\frac{cE_b}{2N_0}\right), \tag{63}$$

where no approximation is necessary. For coherent phase-shift-keyed (PSK), coherent quadriphase-shift-keyed (QPSK), and noncoherent (differential) PSK modulation, we have $c = 2$. For coherent and noncoherent amplitude-shift-keyed (ASK) modulation, we have $c = 1/2$.

Substituting Eq. (61) into both sides of Eq. (60), taking the natural logarithm, and rearranging, we obtain

$$ln\ g(n, k) - \frac{cE_b}{2N_0}\left(\frac{E_b'}{E_b} - 1\right) = \frac{1}{2}\ ln\ \frac{E_b'}{E_b}. \tag{64}$$

We now approximate the right-hand side by the first term in a Taylor-series expansion; that is, we use

$$ln\ \frac{E_b'}{E_b} = ln\left[1 + \left(\frac{E_b'}{E_b} - 1\right)\right] \approx \frac{E_b'}{E_b} - 1, \tag{65}$$

which is reasonably accurate if

$$\frac{E_b'}{E_b} < 1.5. \tag{66}$$

Substituting Eq. (65) into Eq. (64), solving for E_b'/E_b, and employing the result in Eq. (57), we obtain

$$D_C = 10 \log_{10}\left[1 + \frac{2\ ln\ g(n,k)}{\frac{cE_b}{N_0} + 1}\right], \tag{67}$$

where the subscript C is a reminder that this formula holds for coherent modulation systems. Using our solution of Eq. (64) in Eq. (66), the condition for the accuracy of Eq. (67) becomes

$$\frac{E_b}{N_0} > \frac{4 \ln g(n, k) - 1}{c} \tag{68}$$

For noncoherent modulation systems, we obtain in a similar manner

$$D_N = 10 \log_{10} \left[1 + \frac{2 \ln g(n, k)}{\frac{cE_b}{N_0}} \right]. \tag{69}$$

Equation (69) is exact, since neither Eq. (62) nor Eq. (68) is required to derive it. The expressions for D_C and D_N and Eqs. (61) and (63) indicate that, for a fixed plaintext word-error-rate, the degradation is a function of coherency rather than specific modulation type. In other words, the three basic types of coherent systems have the same degradation, and the two basic types of noncoherent systems have the same degradation.

The degradation equations facilitate comparison between block and stream ciphers. An important observation is that for most practical values of n and P_b, the ensemble-average bit-error-rate of block ciphers is nearly the same as that of stream ciphers with $X = 1/2$.

To illustrate some other aspects of block and stream ciphers, an example of noncoherent system degradation shall be studied. Combining Eqs. (40), (63), and (69), we have

$$D_N = 10 \log_{10} \left[1 - \frac{\ln g(n, k)}{\ln \frac{2P_w}{k}} \right]. \tag{70}$$

Figures 4 and 5 are plots of this equation with respect to bit and word ensemble-average error rates when $n = 50$. In Fig. 4 we set $k = 1$ and $P_w = P_b$, and plot D_N as a function of P_b. The function $g(n, k)$ is determined by Eq. (54) for block ciphers and by Eq. (39) for stream ciphers. In Fig. 5 we set $k = 10$ and plot D_N as a function of P_w. The function $g(n, k)$ is determined by Eqs. (40) and (55) for block ciphers and Eqs. (37) and (40) for stream ciphers. It is seen that stream ciphers with $X = 3/4$ cause somewhat less bit-error-rate degradation than the block ciphers. However, the word-error-rate degradation due to block ciphers is lower than that of stream ciphers with $X = 1/2$ over the range of interest. In Figs. 6 and 7 we see the effects of increasing the parameter n when P_b or P_w is fixed. Since n is a measure of the security of the cryptographic system, it appears that the price paid in degradation for increased security is not exorbitant. An interesting observation is that the ensemble-average word-error-rate degradations of block and stream ciphers converge as n increases.

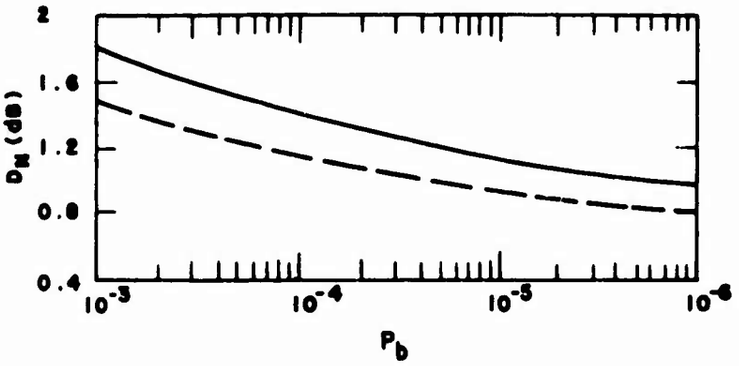


Fig. 4 — Degradation of bit error rate as a function of P_b for a noncoherent system with $n = 50$. Solid curve: block cipher or stream cipher, $X = 1/2$. Dashed curve: stream cipher, $X = 3/4$.

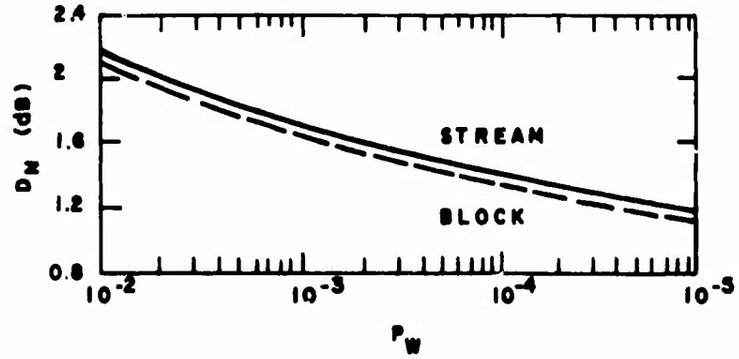


Fig. 5 — Degradation of word error rate as a function of P_w for a noncoherent system with $n = 50$ and $k = 10$.

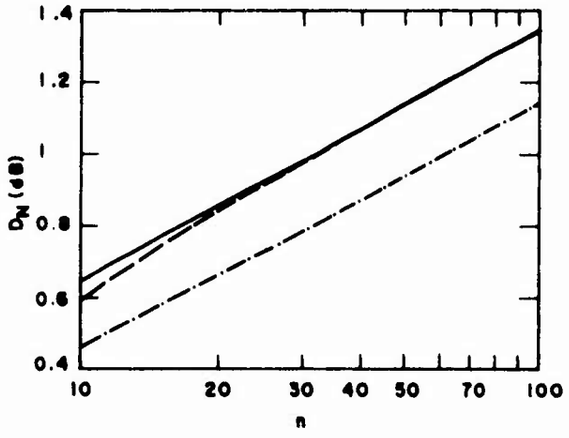


Fig. 6 — Degradation of bit error rate as a function of n for a noncoherent system with $P_b = 10^{-5}$. Solid curve: stream cipher, $X = 1/2$. Dashed curve: block cipher. Dot-dashed curve: stream cipher, $X = 3/4$.

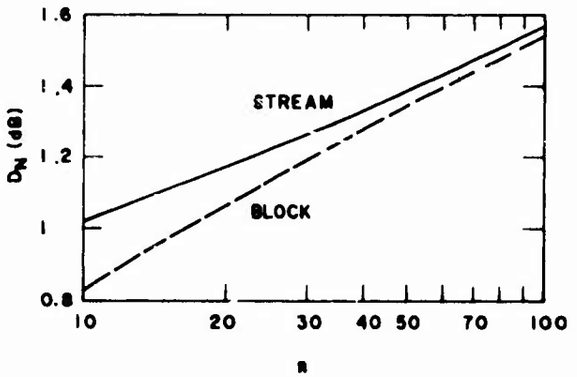


Fig. 7 — Degradation of word error rate as a function of n for a noncoherent system with $P_w = 10^{-4}$ and $k = 10$.

A comparison between Eqs. (48) and (55) reveals that no member of a block-cipher ensemble suffers significantly more word-error-rate degradation than the ensemble average for $n > 3$, $k > 3$, and most practical values of P_b. However, one or more members of a block-cipher ensemble may endure considerably greater bit-error-rate degradation than the ensemble average. For example, with coherent PSK modulation and $n = 60$, it follows from Eq. (67) that some member of the associated block-cipher ensemble may have an extra bit-error-rate degradation ranging from approximately 0.3 dB to 0.2 dB as P_b varies from 10^{-3} to 10^{-6}. Similar statements can be made for stream-cipher ensembles when $X = 1/2$.

Suppose a cryptographic system is provided with the additional power necessary to obtain the same word error rate as the corresponding plaintext system. The question arises as to whether the performance of the cryptographic system is now as good as that of the plaintext system. To answer this question, note that a word error in a plaintext system usually involves one or two erroneous bits. On the other hand, a cryptographic word error usually implies many erroneous bits. Relative performance must be evaluated by determining the additional cost, if any, of multiple bit errors within an erroneous word.

BIBLIOGRAPHY

Perhaps the best introduction to cryptography is provided in the article by Feistel [1]. In addition, Feistel provides an interesting and practical block-cipher scheme. References 2 and 3 contain material concerning the data-keyed cipher of Figs. 1 and 2. Meyer [4] gives a detailed analysis of the vulnerability of linear stream ciphers. A very comprehensive historical exposition with some descriptive technical content is the book by Kahn [5]. One of the rare mathematical treatments of cryptanalysis in English is the book by Sinkov [6]. A study of the mathematical structure of secrecy systems from an information-theory viewpoint is the classic paper of Shannon [7]. The original work on cryptographic error rates, contained in the papers by Torrieri [8,9], has been generalized and extended.

1. H. Feistel, "Cryptography and Computer Privacy," Scientific American 228, No. 5, 15 (1973).

2. T. Twigg, Electronic Design 23, 68 (1972).

3. C.H. Meyer and W.L. Tuchman, Electronic Design 23, 74 (1972).

4. C.H. Meyer, "Enciphered Data for Secure Transmission," Computer Design 13, No. 4, 129 (1974).

5. D. Kahn, *The Codebreakers*, Macmillan, New York, 1967.

6. A. Sinkov, *Elementary Cryptanalysis: A Mathematical Approach*, Random House, New York, 1968.

7. C.E. Shannon, "Communication Theory of Secrecy Systems," Bell System Technical Journal 28, 656 (1949).

8. D.J. Torrieri, "Word Error Rates in Cryptographic Ensembles," IEEE Transactions on Aerospace and Electronic Systems AES-9, 901 (1973).

9. D.J. Torrieri, "Additions and Modifications to 'Word Error Rates in Cryptographic Ensembles'," IEEE Transactions on Aerospace and Electronic Systems AES-10, 715 (1974).